# Chapter Review Exercises (page 795)

R12.1 (a) There is a moderately strong, positive linear relationship between the thickness and the velocity.
(b) $\hat{y} = 70.44 + 274.78x$ where $y$ is the velocity and $x$ is the thickness.
(c) The predicted velocity for $x$ = 0.4 inches is $\hat{y} = 70.44 + 274.78(0.4) = 180.352$ feet/second. The residual is $y - \hat{y}$ = 104.8 – 180.352 = –75.552, so the line overpredicts the velocity by 75.552 ft/sec.
(d) The linear model is appropriate. The scatterplot shows a linear relationship and the residual plot has no leftover patterns.
(e) Slope: For each increase of an inch in thickness, the predicted velocity increases by 274.78 feet/second. $s$: When using the least-squares regression line with $x$ = thickness to predict $y$ = velocity, we will typically be off by about 56.36 feet per second. $r^2$: About 49.3% of the variation in velocity is accounted for by the linear relationship relating velocity to thickness. Standard error of the slope: If we take many different random samples of 12 pistons and compute the least-squares regression line for each sample, the estimated slope will typically vary from the slope of the population regression line for predicting velocity from thickness by about 88.18.

R12.2 *State*: We want to perform a test of $H_0: \beta = 0$ versus $H_a: \beta \neq 0$ where $\beta$ is the slope of the population regression line relating thickness to velocity. We will use $\alpha = 0.05$. *Plan*: If the conditions are met, we will do a $t$ test for the slope $\beta$. Linear: The residual plot shows no leftover patterns, indicating that a linear model is appropriate. Independent: Knowing the velocity for one piston should not help us predict the velocity for another piston. Also, the sample size ($n$ = 12) is less than 10% of the pistons in the population. Normal: We are told that the Normal probability plot of the residuals is roughly linear. Equal SD: The residual plot shows roughly equal scatter for all $x$ values. Random: The data come from a random sample. *Do*: The test statistic is $t = \frac{b-0}{s_b} = \frac{274.78}{88.18} = 3.116$. With df = 12 – 2 = 10, the $P$-value is between 2(0.005) = 0.01 and 2(0.01) = 0.02. *Using technology*: $P$-value = 0.0109.

*Conclude*: Because the $P$-value of 0.0109 is less than $\alpha$ = 0.05 we reject $H_0$. There is convincing evidence of a linear relationship between thickness and gate velocity in the population of pistons formed from this alloy of metal.

R12.3 *State*: We want to construct a 95% confidence interval for $\beta$, the slope of the population regression line relating thickness to velocity. *Plan*: If the conditions are met, we will construct a $t$ interval for the slope $\beta$. The conditions were checked in Exercise R12.2. *Do*: With df = 12 – 2 = 10, the confidence interval is $274.78 \pm 2.228(88.18) = 274.78 \pm 196.465 = (78.315, 471.245)$. *Conclude*: We are 95% confident that the interval from 78.315 to 471.245 captures the slope of the population regression line for predicting velocity from thickness for the population of pistons formed from this alloy of metal. Because 0 is not in the interval, our conclusion is consistent with Exercise R12.2. In both cases, we reject 0 as a plausible value for the slope of the population regression line.

R12.4 The Linear condition is violated because there is clear curvature to the scatterplot and an obvious curved pattern in the residual plot. Also, the Random condition may not be met because we weren't told if the sample was selected at random.

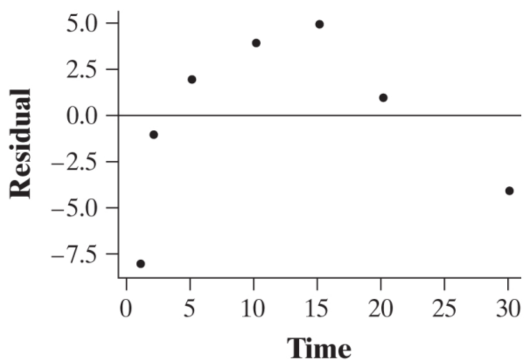
R12.5 (a) The transformation did achieve linearity because there is no leftover pattern in the residual plot.

(b) The equation is $\hat{y} = -0.000595 + 0.3\left(\frac{1}{x^2}\right)$ where $y$ = intensity and $x$ = distance.

(c) For a bulb at a distance of $x$ = 2.1 meters, we would predict an intensity of

$\hat{y} = -0.000595 + 0.3\left(\frac{1}{(2.1)^2}\right) = 0.0674$ candelas.

R12.6 (a) Although answers will vary somewhat, the residual plot should show a negative–positive–negative pattern like the one below.




(b) A linear model is not appropriate because the scatterplot clearly shows that the relationship between practice time and percent of words recalled is curved and the residual plot has a leftover pattern.
(c) A power model is more appropriate in this case because the scatterplot showing ln(recall) versus ln(time) is more linear than the scatterplot showing ln(recall) versus time.
(d) For the power model, $\widehat{\ln y} = 3.48 + 0.293\ln(25) = 4.423$ and $\hat{y} = e^{4.423} = 83.35$ percent of words recalled. For the exponential model, $\widehat{\ln y} = 3.69 + 0.0304(25) = 4.45$ and $\hat{y} = e^{4.45} = 85.63$ percent of words recalled. Based on my answer to part (c), I think the power model will give a better prediction.